

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: Moore et al.

Application Serial No.: Unassigned

Art Unit: Unassigned

Filed: Herewith

Examiner: Unassigned

For: **Human FK506 Binding Proteins** Attorney Docket No.: **PF392**

PATENT APPLICATION FEE VALUE

Assistant Commissioner
for Patents
Box PATENT APPLICATION
Washington, D.C. 20231

Sir:

Transmitted herewith for filing under 37 C.F.R. §1.53 is the following patent application:

Applicants: Paul A. Moore
Ying-Fei Wei
Steven M. Ruben

Executed on: Unexecuted
Unexecuted
Unexecuted

Title of Invention: **Human FK506 Binding Proteins**

Including: Specification (54 pgs); 20 Claims (3 pgs); Abstract (1 pg);
Sequence Listing (17 pgs); 4 Figures (10 sheets)

The filing fee the subject application being filed concurrently herewith has been calculated as follows:


TYPE	NO. FILE	LESS	EXTRA	EXTRA RATE	FEE
Total Claims	20	-20	0	\$18.00 each	0.00
Independent	4	-3	1	\$78.00 each	\$78.00
Minimum Fee					\$760.00
Multiple Dependency Fee if applicable (\$260.00)					
Total Filing Fee					\$838.00

This application claims benefit of 35 U.S.C. section 119(e) based on copending U.S. Provisional Application Serial No. 60/070,875 filed January 9, 1998.

Please charge the required fee to Deposit Account No. 08-3425. In addition, the Commissioner is hereby authorized to charge payment for any additional filing fees required under 37 C.F.R. 1.16 or credit any overpayment to Deposit Account No. 08-3425. A duplicate of this paper is attached.

Respectfully submitted,

Date: 1/06/99


A. Anders Brookes (Reg. No. 36,373)
Attorney for Applicants

Human Genome Sciences, Inc.
9410 Key West Avenue
Rockville, MD 20850
(301) 610-5761 (telephone)

Enclosure
AAB/mbp

JCS11 U.S. PTO
09/225502
01/06/99

09225502 01/06/99

66/90/10
JCS11 U.S. PTO
01/06/99

Please type a plus sign (+) inside this box →



PTO
01/06/99
05/25502

UTILITY PATENT APPLICATION TRANSMITTAL <small>(Only for new nonprovisional applications under 37 CFR 1.53(b))</small>		Attny Docket No. PF392 Total Pages: 87 First Named Inventor or Application Identifier Paul A. Moore Date January 6, 1999
APPLICATION ELEMENTS <small>See MPEP chapter 600 concerning utility patent application contents</small>		Assistant Commissioner for Patents Address to: Box Patent Application Washington, D.C. 20231
<div style="display: flex; justify-content: space-between;"> <div style="width: 48%;"> <p>1. <input checked="" type="checkbox"/> Fee Transmittal Form</p> <p>2. <input checked="" type="checkbox"/> Specification [total pgs. 58 (preferred arrangement set forth below) - Descriptive title of the Invention - Cross References to Related Applications - Statement Regarding Fed Sponsored R&D - Reference to Microfiche Appendix - Background of the Invention - Brief Summary of the Invention - Brief Description of the Drawings (if filed) - Detailed Description - Claim(s) - Abstract of the Disclosure</p> <p>3. <input checked="" type="checkbox"/> Drawing(s) 4 Total Sheets 10</p> <p>4. <input checked="" type="checkbox"/> Oath or Declaration Total pages 2 a. <input type="checkbox"/> Newly executed (original or copy) b. <input type="checkbox"/> Copy from a prior application (1.63(d)) (for con/div with Box 17 completed) i. <input type="checkbox"/> Deletion of Inventors (signed statement attached deleting inventor(s) named in prior app.)</p> <p>5. <input type="checkbox"/> Incorporation by Reference (with Box 4b) The entire disclosure of the prior application, from which a copy of the oath or declaration is supplied under Box 4b, is considered as being part of the disclosure of the accompanying application and is hereby incorporated by reference therein.</p> </div> <div style="width: 48%;"> <p>6. <input type="checkbox"/> Microfiche Computer Program (appendix)</p> <p>7. Nucleotide and/or Amino Acid Sequence Submission (if applicable, all necessary) a. <input checked="" type="checkbox"/> Computer Readable Copy b. <input checked="" type="checkbox"/> Paper Copy (identical to computer copy) [17pgs] c. <input checked="" type="checkbox"/> Statement verifying identity of above copies</p> </div> </div>		
		Accompanying Application Parts 8. <input type="checkbox"/> Assignment Papers (cover and document(s)) 9. <input type="checkbox"/> 37 CFR 3.73(b) Statement <input type="checkbox"/> Power of Attny 10. <input type="checkbox"/> English Translation Document 11. <input type="checkbox"/> Information Disclosure Statement/PTO-1449 <input type="checkbox"/> Copies of IDS Citations 12. <input type="checkbox"/> Preliminary Amendment 13. <input checked="" type="checkbox"/> Return Receipt Postcard (MPEP 503) (should be specifically itemized) 14. <input type="checkbox"/> Small Entity Statement(s) <input type="checkbox"/> Statement filed in prior application, status still proper and desired 15. <input type="checkbox"/> Certified copy of priority document(s) 16. <input type="checkbox"/> Other:
17. If a CONTINUING APPLICATION, check appropriate box and supply the requisite information: <input type="checkbox"/> Continuation <input type="checkbox"/> Divisional <input type="checkbox"/> Continuation-in-Part (CIP) of prior application No.:		
18. CORRESPONDENCE ADDRESS		
<input type="checkbox"/> Customer Number or Bar Code Label or <input checked="" type="checkbox"/> Correspondence address below		
Name	A. Anders Brookes, Esq., Attorney for Applicants Human Genome Sciences, Inc.	
Address	9410 Key West Avenue	
City	Rockville	State MD Zip Code 20850
Country	US	Telephone 301-309-8504 Fax 301-309-8439
 Signature		Reg. No. 36,373 1/06/99 Date

09235502, 010699

Human FK506 Binding Proteins

This application claims priority under 35 U.S.C. §119(e) to copending US Provisional Patent Application Serial No. 60/070,875 filed January 9, 1998, hereby incorporated herein by reference.

Field of the Invention

This invention relates to newly identified human polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and their production. More particularly the invention provides novel FK506 Binding Protein polypeptides and polynucleotides encoding such polypeptides. Also provided are methods of using such novel polynucleotides and polypeptides.

Background of the Invention

FK506 and rapamycin are structurally related macrolides that block distinct steps in intracellular signalling pathways. (Sawada, S. et al., J. Immunol., 139:1797-1803 (1987); Tocci, M. J., et al., J. Immunol., 143:618-726/(1989)). Both are potent immunosuppressants, and drug action is mediated in part by binding to members of the immunophilin protein family. (Schreiber, S. L., Science, 251:283-287 (1991); Rosen, M. K. and Schrieber, S. L., Angew. Chem. Int. Ed. Engl., 31:384-400 (1992)). One recently identified FK506 binding protein (FKBP) is FKBP12 with approximate relative molecular mass (M.sub.r) of 11,800 (12K), and a PI of 8.8-8.9. (Harding, M. W., et al., Nature, 341:758-760 (1989)). Studies have shown that the unbound FKBP12 catalyzes the cis-trans isomerization of proline residues in proteins and peptides. However, when FKBP12 binds FK506, this activity is inhibited. Recent studies suggest that the FK506-FKBP12 complex functions as an immunosuppressant by binding to, and altering, the phosphatase activity of calcineurin/calmodulin.

Thus, there is a need for identification and characterization of novel FK506 Binding Protein polypeptides that have the ability to modulate immune system function.

Summary of the Invention

The present invention relates to novel FK506 Binding Protein polypeptides and to polynucleotides that encode such polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

Brief Description of the Drawings

Figure 1 shows the nucleotide sequence (SEQ ID NO:1) of the human cDNA in clone HMEAA94 and the deduced amino acid sequence (SEQ ID NO:2) encoded thereby.

5 Figure 2 shows the nucleotide sequence (SEQ ID NO:3) of the human cDNA in clone HL1AP03 and the deduced amino acid sequence (SEQ ID NO:4) encoded thereby.

Figure 3 shows the nucleotide sequence (SEQ ID NO:5) of the human cDNA in clone HSYBM46 and the deduced amino acid sequence (SEQ ID NO:6) encoded thereby.

Figure 4 shows the nucleotide sequence (SEQ ID NO:7) of the human cDNA in clone HFKBC47 and the deduced amino acid sequence (SEQ ID NO:8) encoded thereby.

10

Detailed Description

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

15 In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide. "Isolated", however, is not intended to include complete and intact human chromosomes whether contained in a cell or removed from a cell.

20 In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

25 As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

30 A clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801

09235509 010699

University Boulevard, Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC.

"Stringent hybridization conditions" refers to an overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH₂PO₄; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 ug/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and

RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, *PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES*, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); *POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS*, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., *Meth Enzymol* 182:626-646 (1990); Rattan et al., *Ann NY Acad Sci* 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the

polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

Polynucleotides and Polypeptides of the Invention

FEATURES OF PROTEIN ENCODED BY SEQ ID NO: 1 (GENE 1)

The translation product of this gene (SEQ ID NO:2) shows very strong homology to FK506 Binding Protein 65 ("FKBP65"). See Coss M.C. et al., J. Bio. Chem., (1995) 270:29336 (GenBank accession no. g894162). FKBP65 is a member of the immunophilin superfamily. FK506 binding proteins are known to bind immune modulating drugs FK506, rapamycin, and cyclosporin A. Without expression of the FK506 binding protein the aforementioned drugs do not have activity. Thus, FK506 binding proteins play an essential role in effecting the activities of these immunosuppressant drugs.

The cDNA clone (HMEAA94) deposited with the ATCC herein and used to generate SEQ ID NO:1 was obtained from a cDNA library made from microvascular endothelial cells.

The homology to FKBP65 suggests that the protein product of this clone would be useful in screening assays for the discovery of novel immunosuppressant drugs and in the treatment of diseases caused by an over-active immune system, particularly those caused by T-cells, such as graft vs. host disease, rheumatoid arthritis, inflammation and osteoarthritis. See, for example, US Patent No. 5,498,597, which describes uses for the novel FKBP-13 polypeptide. The uses described in US Patent No. 5, 498,597, hereby incorporated by reference in its entirety, are equally applicable to the translation products of SEQ ID NO:1.

Immunophilins possess peptidylproline cis-trans-isomerase (PPIase) activity. An assay for measuring such activity is described in Galat, A., Eur. J. Biochem., (1993) 216:689, incorporated herein by reference in its entirety. Polypeptides of the invention preferably include those having PPIase activity.

FEATURES OF PROTEIN ENCODED BY SEQ ID NO: 3 (GENE 2)

SEQ ID NO:3 shows the nucleic acid sequence of a partial cDNA which is a splice variant of the gene described as SEQ ID NO:1. The translation product of this gene (SEQ ID NO:4) also shows very strong homology to FK506 Binding Protein 65 ("FKBP65"). See Coss M.C. et al., J. Bio. Chem., (1995) 270:29336.

As mentioned above, FKBP65 is a member of the immunophilin superfamily. FK506 binding proteins are known to bind immune modulating drugs FK506, rapamycin, and cyclosporin A. Without expression of the FK506 binding protein the aforementioned drugs do

not have activity. Thus, FK506 binding proteins play an essential role in effecting the activities of these immunosuppressant drugs.

The homology to FKBP65 suggests that the protein product of this clone would be useful for in screening assays for the discovery of novel immunosuppressant drugs and in the treatment of diseases caused by over-active immune system, particularly those caused by T-cells, such as graft vs. host disease, rheumatoid arthritis, inflammation and osteoarthritis. See, for example, US Patent No. 5,498,597, which describes uses for the novel FKBP-13 polypeptide. The uses described in US Patent No. 5, 498,597, hereby incorporated by reference in its entirety, are equally applicable to the translation products of SEQ ID NO:3.

Immunophilins possess peptidylproline cis-trans-isomerase (PPIase) activity. An assay for measuring such activity is described in Galat, A., Eur. J. Biochem., (1993) 216:689, incorporated herein by reference in its entirety. Polypeptides of the invention preferably include those having PPIase activity.

A full-length cDNA clone, and corresponding protein, may be produced by those of ordinary skill in the art by aligning SEQ ID NO:1 and SEQ ID NO:3. Since these clones are nearly identical until they diverge at the 3' end, the 5' portion of SEQ ID NO:3 may be supplied by using the coding region from SEQ ID NO:1.

FEATURES OF PROTEIN ENCODED BY SEQ ID NO: 5 (GENE 3)

The translation product of this gene (SEQ ID NO:6) shows very strong homology to FK506 Binding Protein 65 ("FKBP65"). See Coss M.C. et al., J. Bio. Chem., (1995) 270:29336. FKBP65 is a member of the immunophilin superfamily. FK506 binding proteins are known to bind immune modulating drugs FK506, rapamycin, and cyclosporin A. Without expression of the FK506 binding protein the aforementioned drugs do not have activity. Thus, FK506 binding proteins play an essential role in effecting the activities of these immunosuppressant drugs.

The cDNA clone (HSYBM46) deposited with the ATCC herein and used to generate SEQ ID NO:5 was obtained from a cDNA library made from thymus stromal cells.

The homology to FKBP65 suggests that the protein product of this clone would be useful for in screening assays for the discovery of novel immunosuppressant drugs and in the treatment of diseases caused by over-active immune system, particularly those caused by T-cells, such as graft vs. host disease, rheumatoid arthritis, inflammation and osteoarthritis. See, for example, US Patent No. 5,498,597, which describes uses for the novel FKBP-13 polypeptide. The uses described in US Patent No. 5, 498,597, hereby incorporated by reference in its entirety, are equally applicable to the translation product of SEQ ID NO:5.

Immunophilins possess peptidylproline cis-trans-isomerase (PPIase) activity. An assay for measuring such activity is described in Galat, A., Eur. J. Biochem., (1993) 216:689, incorporated herein by reference in its entirety. Polypeptides of the invention preferably include those having PPIase activity.

0925503-010699

FEATURES OF PROTEIN ENCODED BY SEQ ID NO: 7 (GENE 4)

SEQ ID NO:4 shows the nucleic acid sequence of a cDNA which is a splice variant of the gene described as SEQ ID NO:5. The translation product of this gene (SEQ ID NO:8) also shows very strong homology to FK506 Binding Protein 65 ("FKBP65"). See Coss M.C. et al., J. Bio. Chem., (1995) 270:29336.

As mentioned above, FKBP65 is a member of the immunophilin superfamily. FK506 binding proteins are known to bind immune modulating drugs FK506, rapamycin, and cyclosporin A. Without expression of the FK506 binding protein the aforementioned drugs do not have activity. Thus, FK506 binding proteins play an essential role in effecting the activities of these immunosuppressant drugs.

The cDNA clone (HFKBC47) deposited with the ATCC herein and used to generate SEQ ID NO:7 was obtained from a cDNA library made from fetal kidney tissue.

The homology to FKBP65 suggests that the protein product of this clone would be useful for in screening assays for the discovery of novel immunosuppressant drugs and in the treatment of diseases caused by over-active immune system, particularly those caused by T-cells, such as graft vs. host disease, rheumatoid arthritis, inflammation and osteoarthritis. See, for example, US Patent No. 5,498,597, which describes uses for the novel FKBP-13 polypeptide. The uses described in US Patent No. 5, 498,597, hereby incorporated by reference in its entirety, are equally applicable to the translation products of SEQ ID NO:7.

Immunophilins possess peptidylproline cis-trans-isomerase (PPIase) activity. An assay for measuring such activity is described in Galat, A., Eur. J. Biochem., (1993) 216:689, incorporated herein by reference in its entirety. Polypeptides of the invention preferably include those having PPIase activity.

A full-length cDNA clone, and corresponding protein, may be produced by those of ordinary skill in the art by aligning SEQ ID NO:3 and SEQ ID NO:4. Since these clone are nearly identical until they diverge at the 3' end, the 5' portion of SEQ ID NO:4 may be supplied by using the coding region from SEQ ID NO:3.

09235502 "010699

TABLE 1. FEATURES OF PARTICULAR SECRETED PROTEINS

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	N O T E	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
1	HMEAA94	209193 08/01/97	pBS SK-	1	1234	-60	1174		1	1	2	1	26	27	336
2	HL1AP03	209193 08/01/97	pBS SK-	3	2145	1	2145		2		4			1	441
3	HSYBM46	209193 08/01/97	pCMVSPORT 3.0	5	3391	1	3391		130	130	6	1	24	25	574
4	HFKBC47	209193 08/01/97	pBS SK-	7	1251	1	1251		3		8			1	388

Table 1, above, summarizes the information corresponding to each "Gene No." described above.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." "Vector" refers to the type of vector into which the human cDNA has been inserted.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the proteins encoded by the cDNA clones identified in Table 1.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention

09225502 "010699

deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits.

Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below). It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies of the invention raised against the protein in methods which are well known in the art.

Signal Sequences

Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of

these methods is in the range of 75-80%. (von Heinje, supra.) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

In the present case, the deduced amino acid sequence of the secreted polypeptide was analyzed by a computer program called SignalP (Henrik Nielsen et al., Protein Engineering 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results shown in Table 1.

As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., + or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Polynucleotide and Polypeptide Variants

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence shown in Table 1, the ORF (open reading frame), or any fragment specified as described herein.

As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the present invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in Table 1 or to the amino acid sequence encoded by deposited DNA clone can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C- termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host such as *E. coli*).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the protein without substantial loss of biological function. The authors of Ron et al., *J. Biol. Chem.* 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., *J. Biotechnology* 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (*J. Biol. Chem* 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that

averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See, Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as to have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., *Science* 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, *Science* 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino

acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

Polynucleotide and Polypeptide Fragments

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably,

these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the complete protein as well as the mature form. Further preferred polypeptide fragments include the complete protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can

bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998-4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the complete protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')₂ fragments) which are capable of specifically binding to protein. Fab and F(ab')₂ fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

09225502, 010699

Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986).

It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid
 5 extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be
 10 recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention
 15 may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is
 20 efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents.
 25 The following description should be considered exemplary and utilizes known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome
 marking reagents, based on actual sequence data (repeat polymorphisms), are presently available.

Each polynucleotide of the present invention can be used as a chromosome marker.
 30

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be selected
 using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing
 35 individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal
 cyclor. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific

chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library) .) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano,

J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once a unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify

0923502 " 010699

tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

10 Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell . Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (125I, 121I), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and technetium (99mTc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, 131I, 112In, 99mTc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of 99mTc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled

Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

Biological Activities

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

Immune Activity

A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and

macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or
 5 detector of a particular immune system disease or disorder.

A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a
 10 decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs),
 15 Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia,
 20 factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of
 25 a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis,
 35 Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitus, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

5 A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the
10 present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the
15 proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung
20 injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

Infectious Disease

A polypeptide or polynucleotide of the present invention can be used to treat or detect
25 infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune
30 response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae,
35 Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus).

Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiolitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox ,

5 hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be

10 treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses,

15 Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Meningococcal), Pasteurellaceae Infections (e.g., Actinobacillus, Haemophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but

20 not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy,

25 Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or

30 symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

35

Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Binding Activity

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

09225502.010699

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity, and (b) determining if a biological activity of the polypeptide has been altered.

Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

US 2010/010699

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

5 A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

10 A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

15 Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

20 Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

25 Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

30 Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

35 A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a

nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the complete protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid
 5 sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type of a
 10 biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated
 15 with abnormal structure or expression of a gene encoding a protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at
 20 least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

In any of these methods, the step of detecting said polypeptide molecules includes using
 25 an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at
 30 least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence
 35 encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a

0925502-010699

protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

Examples

Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector. Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

Vector Used to Construct Library

Corresponding Deposited Plasmid

Lambda Zap

pBluescript (pBS)

Uni-Zap XR

pBluescript (pBS)

Zap Express	pBK
lafmid BA	plafmid BA
pSport1	pSport1
pCMVSPORT 2.0	pCMVSPORT 2.0
pCMVSPORT 3.0	pCMVSPORT 3.0
pCR [®] 2.1	pCR [®] 2.1

Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988);

Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS-. The S and K refers to the orientation of the polylinker to the T7 and T3 primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the other, antisense.

Vectors pSport1, pCMVSPORT 2.0 and pCMVSPORT 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into E. coli strain XL-1 Blue. Vector pCR[®]2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al., Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with ^{32}P - γ -ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 μl of reaction mixture with 0.5 μg of the above cDNA template. A convenient reaction mixture is 1.5-5 mM MgCl_2 , 0.01% (w/v) gelatin, 20 μM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., *Nucleic Acids Res.* 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the

gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide

A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

Example 3: Tissue Distribution of Polypeptide

Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P³² using the rediprime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed according to standard procedures.

Example 4: Chromosomal Mapping of the Polynucleotides

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This primer set is then used in a

polymerase chain reaction under the following set of conditions : 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is
 5 analyzed on either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

Example 5: Bacterial Expression of a Polypeptide

10 A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and
 15 XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated
 20 into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^r). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed
 25 by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.⁶⁰⁰) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is
 30 then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant
 35 containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA

resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM imidazole. Imidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains: 1) a neomycinphosphotransferase gene as a selection marker, 2) an E. coli origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (lacIq). The origin of replication (oriC) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

Example 6: Purification of a Polypeptide from an Inclusion Body

The following alternative method can be used to purify a polypeptide expressed in *E. coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A_{280} monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 µg of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

Example 7: Cloning and Expression of a Polypeptide in a Baculovirus

Expression System

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak *Drosophila* promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., *Virology* 170:31-39 (1989).

Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("GeneClean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("GeneClean" BIO 101 Inc., La Jolla, Ca.).

The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and
 5 analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

Five µg of a plasmid containing the polynucleotide is co-transfected with 1.0 µg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., Proc.
 10 Natl. Acad. Sci. USA 84:7413-7417 (1987). One µg of BaculoGold™ virus DNA and 5 µg of the plasmid are mixed in a sterile well of a microtiter plate containing 50 µl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 µl Lipofectin plus 90 µl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm
 15 tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described
 20 by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques
 25 are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

30 To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc.,
 35 Rockville, MD). After 42 hours, 5 µCi of ³⁵S-methionine and 5 µCi ³⁵S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

Example 8: Expression of a Polypeptide in Mammalian Cells

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLV, HIV and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSport 2.0, and pCMVSport 3.0. Mammalian host cells that could be used include, human HeLa, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No. 209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites

BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("GeneClean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five μ g of the expression plasmid pC6 is cotransfected with 0.5 μ g of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 μ M, 2 μ M, 5 μ M, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 μ M. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

Example 9: Protein Fusions

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394,827; Traunecker, et al., Nature 331:84-86

09225502 "010699

(1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the half-life time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

Human IgG Fc region:

GGGATCCGGAGCCCAAATCTTCTGACAAACTCACACATGCCCACCGTGCCCAGCAC
CTGAATTCGAGGGTGCACCGTCAGTCTTCTCTTCCCCCAAACCCAAGGACACCC
TCATGATCTCCCGGACTCCTGAGGTCACATGCGTGGTGGTGGACGTAAGCCACGAA
GACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAA
GACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCA
CCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAAC
AAAGCCCTCCCAACCCCCATCGAGAAAACCATCTCCAAGCCAAAGGGCAGCCCCG
AGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGG
TCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGTGGAGTGG
GAGAGCAATGGGCAGCCGGAGAACAACACTACAAGACCACGCCTCCCGTGCTGGACTC
CGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGC
AGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGC
AGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAGTGCGACGGCCGCGACTCTAGAGG
AT (SEQ ID NO:13)

Example 10: Production of an Antibody from a Polypeptide

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., Nature 256:495 (1975); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (Gastroenterology 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

It will be appreciated that Fab and F(ab')₂ and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab

fragments) or pepsin (to produce F(ab')₂ fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

Example 11: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies). The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenindeoxy-uridine 5'-triphosphate (Boehringer Mannheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv. et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 12: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 10. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

Example 13: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of

delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 $\mu\text{g/kg/day}$ to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 $\mu\text{g/kg/hour}$ to about 50 $\mu\text{g/kg/hour}$, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22:547-556 (1983)), poly (2- hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and R. Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic

to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological

products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

5 **Example 14: Method of Treating Decreased Levels of the Polypeptide**

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

20 **Example 15: Method of Treating Increased Levels of the Polypeptide**

Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

25 **Example 16: Method of Treatment Using Gene Therapy**

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

5 The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is
10 maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS),
15 penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the
20 infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a
25 retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

It will be clear that the invention may be practiced otherwise than as particularly described
30 in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of
35 the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both incorporated herein by reference in their entireties.

09225502 "010699

What Is Claimed Is:

1. An isolated nucleic acid molecule comprising a nucleotide sequence encoding an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:

- (a) residues 27 to 336 of SEQ ID NO:2;
- (b) residues 1 to 441 of SEQ ID NO:4;
- (c) residues 25 to 574 of SEQ ID NO:6; and
- (d) residues 1 to 388 of SEQ ID NO:8.

2. The isolated nucleic acid molecule of claim 1 comprising the nucleotide sequence shown as nucleotides 1-1008 of SEQ ID NO:1.

3. The isolated nucleic acid molecule of claim 1 comprising the nucleotide sequence shown as nucleotides 2-1324 of SEQ ID NO:3.

4. The isolated nucleic acid molecule of claim 1 comprising the nucleotide sequence shown as nucleotides 130-1751 of SEQ ID NO:5.

5. The isolated nucleic acid molecule of claim 1 comprising the nucleotide sequence shown as nucleotides 3-1166.

6. An isolated nucleic acid molecule comprising a nucleotide sequence encoding an amino acid sequence selected from the group consisting of:

- (a) at least 30 contiguous amino acid residues of SEQ ID NO:2;
- (b) at least 30 contiguous amino acid residues of SEQ ID NO:4;
- (c) at least 30 contiguous amino acid residues of SEQ ID NO:6; and
- (d) at least 30 contiguous amino acid residues of SEQ ID NO:8.

7. A vector comprising the isolated nucleic acid molecule of claim 1.

8. A nucleic acid molecule comprising the nucleic acid molecule of claim 1 operably associated with a heterologous regulatory element which controls gene expression.

9. A host cell comprising the vector or claim 7.

10. A host cell comprising the nucleic acid molecule of claim 8.

11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:

09235502-010599

- (a) residues 27 to 336 of SEQ ID NO:2;
- (b) residues 1 to 441 of SEQ ID NO:4;
- (c) residues 25 to 574 of SEQ ID NO:6; and
- (d) residues 1 to 388 of SEQ ID NO:8.

5

12. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

- (a) at least 30 contiguous amino acid residues of SEQ ID NO:2;
- (b) at least 30 contiguous amino acid residues of SEQ ID NO:4;
- (c) at least 30 contiguous amino acid residues of SEQ ID NO:6; and
- (d) at least 30 contiguous amino acid residues of SEQ ID NO:8.

10

13. An isolated antibody that binds specifically to the isolated polypeptide of claim

11.

15

14. A composition comprising the polypeptide of claim 11.

15. A method of making an isolated polypeptide comprising:

- (a) culturing the host cell of claim 10 under conditions such that said polypeptide is expressed; and
- (b) recovering said polypeptide.

20

16. The polypeptide produced by the method of claim 15.

25

17. A method for treating a medical condition, comprising administering to a patient a therapeutically effective amount of the polypeptide of claim 11.

18. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

30

- (a) determining the presence or absence of a mutation in the polynucleotide of claim 1; and
- (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.

35

19. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

- (a) determining the presence or amount of expression of the polypeptide of claim 11 in a biological sample; and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

- 5 20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:
- (a) contacting the polypeptide of claim 11 with a binding partner; and
 - (b) determining whether the binding partner effects an activity of the polypeptide.

09235503 01069
669010 20552260

Abstract

5 The present invention relates to novel human FK506 binding proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human FK506 binding proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human proteins.

09295502 010599

FIGURE 1A

HMEAA94

```

-50              -30              -10
GGAGGTTGGTGGCGACTCCCTCGCTCGCCCTCACTGCCGGCGGTCCCAACTCCAGGCACC
   10              30              50
ATGTTCCCGCGGGCCCCCCCCAGCCACAGCCTCCTCCGGCTCCCCCTGCTGCAGTTGCTG
M F P A G P P S H S L L R L P L L Q L L
   70              90             110
CTACTGGTGGTGCAGGCCGTGGGGAGGGGGCTGGGCCGCGCCAGCCCGCGGGGGCCCC
L L V V O A V G R G L G R A S P A G G P
   130             150             170
CTGGAAGATGTGGTCATCGAGAGGTACCACATCCCCAGGGCCTGTCCCCGGGAAGTGCAG
L E D V V I E R Y H I P R A C P R E V Q
   190             210             230
ATGGGGGATTTTGTGCGCTACCACTACAACGGCACTTTTGAAGATGGCAAGAAGTTTGAT
M G D F V R Y H Y N G T F E D G K K F D
   250             270             290
TCAAGCTATGATCGCAACACCTTGGTGGCCATCGTGGTGGGTGTGGGGCGCCTCATCACT
S S Y D R N T L V A I V V G V G R L I T
   310             330             350
GGCATGGACCGAGGCCTCATGGGCATGTGTGCAACGAGCGGCGACGCCTCATTGTGCCT
G M D R G L M G M C V N E R R R L I V P
   370             390             410
CCCCACCTGGGCTATGGGAGCATCGGCCTGGCGGGGCTCATTCCACCGGATGCCACCCTC
P H L G Y G S I G L A G L I P P D A T L
   430             450             470
TACTTCGATGTGGTTCTGCTGGATGTGTGGAACAAGGAAGACACCGTGCAGGTGAGCACA
Y F D V V L L D V W N K E D T V Q V S T
   490             510             530
TTGCTGCGCCCGCCCCACTGCCCCCGCATGGTCCAGGACGGCGACTTTGTCCGCTACCAC
L L R P P H C P R M V Q D G D F V R Y H
   550             570             590
TACAATGGCACCCTGCTGGACGGCACCTCCTTCGACACCAGCTACAGTAAGGGCGGCACT
Y N G T L L D G T S F D T S Y S K G G T
   610             630             650
TATGACACCTACGTCGGCTCTGCTTGGCTGATCAAGGGCATGGACCAGGGGCTGCTGGGC
Y D T Y V G S G W L I K G M D Q G L L G
   670             690             710
ATGTGTCCTGGACAGAGAAGGAAGATTATCATCCCTCCATTCTGGCCTATGGCGAGAAA
M C P G Q R R K I I I P P F L A Y G E K
   730             750             770
GGCTATGGTGAAGGTGGGCAAGGACACAAGGGGAAATTCCGCAGAAGAGGGAAAAACCAG
G Y G E G G Q G H K G K F R R R G K N Q

```

09225502-010699

2/10

FIGURE 1B

790 810 830
GCCTCCACATACAGTTGCTCAGGTTGTATACTGCACGAGGGCATCCAACCAAGGACTCAA
A S T Y S C S G C I L H E G I Q P R T Q
850 870 890
GGTGGGATGAAATCTACCCTTGGTGCTACTAAGAAGGGGTGCTTTGGCCGGGCGTGGTGG
G G M K S T L G A T K K G C F G R A W W
910 930 950
CTCACGCTTGTAATCCCAGCACTTTGGGAAGCCAAGGCGGGAGGATCACGAGGTCCAGGA
L T L V I P A L W E A K A G G S R G P G
970 990 1010
GATCGAGACCACGGTGAAACCCCGTCTCTACTAAAAATACAAAAAATTAGCCGGGCGTG
D R D H G E T P S L L K I Q K N *
1030 1050 1070
GTGGGGGCGCCTGTAGTCCCAGCTACTCGGAGAGGCTGAGGCAGGAAAATGACGTGAACC
1090 1110 1130
CGGGAGGCGGAGCTTGCAGTGAGCCGAGATCTTGCCACTGCACTCCAGCCTGGGTGACAG
1150 1170
AGCGAGACTCTGTCTCAAAAAAAAAAAAAAAAAA

09225502.010699

SEQUENCE LISTING

<110> Moore, et al.

<120> Human FK506 Binding Proteins

<130> PF392

<140> Unassigned

<141> 1999-01-06

<150> 60/070,875

<151> 1998-01-09

<160> 8

<170> PatentIn Ver. 2.0

<210> 1

<211> 1234

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (61)..(1008)

<400> 1

ggaggttggt ggcgactccc tcgctcgccc tcaactgccg cggtcccaac tccaggcacc 60

atg	ttc	ccc	gcg	ggc	ccc	ccc	agc	cac	agc	ctc	ctc	cgg	ctc	ccc	ctg	108
Met	Phe	Pro	Ala	Gly	Pro	Pro	Ser	His	Ser	Leu	Leu	Arg	Leu	Pro	Leu	
1				5					10					15		

ctg	cag	ttg	ctg	cta	ctg	gtg	gtg	cag	gcc	gtg	ggg	agg	ggg	ctg	ggc	156
Leu	Gln	Leu	Leu	Leu	Leu	Val	Val	Gln	Ala	Val	Gly	Arg	Gly	Leu	Gly	
			20					25					30			

cgc	gcc	agc	cgc	gcc	ggg	ggc	ccc	ctg	gaa	gat	gtg	gtc	atc	gag	agg	204
Arg	Ala	Ser	Pro	Ala	Gly	Gly	Pro	Leu	Glu	Asp	Val	Val	Ile	Glu	Arg	
		35					40					45				

tac	cac	atc	ccc	agg	gcc	tgt	ccc	cgg	gaa	gtg	cag	atg	ggg	gat	ttt	252
Tyr	His	Ile	Pro	Arg	Ala	Cys	Pro	Arg	Glu	Val	Gln	Met	Gly	Asp	Phe	
	50					55					60					

gtg	cgc	tac	cac	tac	aac	ggc	act	ttt	gaa	gat	ggc	aag	aag	ttt	gat	300
Val	Arg	Tyr	His	Tyr	Asn	Gly	Thr	Phe	Glu	Asp	Gly	Lys	Lys	Phe	Asp	
65					70				75					80		

tca	agc	tat	gat	cgc	aac	acc	ttg	gtg	gcc	atc	gtg	gtg	ggg	gtg	ggg	348
Ser	Ser	Tyr	Asp	Arg	Asn	Thr	Leu	Val	Ala	Ile	Val	Val	Gly	Val	Gly	
				85					90					95		

cgc	ctc	atc	act	ggc	atg	gac	cga	ggc	ctc	atg	ggc	atg	tgt	gtc	aac	396
Arg	Leu	Ile	Thr	Gly	Met	Asp	Arg	Gly	Leu	Met	Gly	Met	Cys	Val	Asn	
			100					105					110			

gag	cgg	cga	cgc	ctc	att	gtg	cct	ccc	cac	ctg	ggc	tat	ggg	agc	atc	444
Glu	Arg	Arg	Arg	Leu	Ile	Val	Pro	Pro	His	Leu	Gly	Tyr	Gly	Ser	Ile	
		115					120					125				

0935502 010699

```
<210> 2
<211> 316
<212> PRT
<213> Homo sapiens
```


Met	Phe	Pro	Ala	Gly	Pro	Pro	Ser	His	Ser	Leu	Leu	Arg	Leu	Pro	Leu
1				5					10					15	
Leu	Gln	Leu	Leu	Leu	Leu	Val	Val	Gln	Ala	Val	Gly	Arg	Gly	Leu	Gly
			20					25					30		
Arg	Ala	Ser	Pro	Ala	Gly	Gly	Pro	Leu	Glu	Asp	Val	Val	Ile	Glu	Arg
		35					40					45			
Tyr	His	Ile	Pro	Arg	Ala	Cys	Pro	Arg	Glu	Val	Gln	Met	Gly	Asp	Phe
	50					55					60				
Val	Arg	Tyr	His	Tyr	Asn	Gly	Thr	Phe	Glu	Asp	Gly	Lys	Lys	Phe	Asp
65					70					75					80
Ser	Ser	Tyr	Asp	Arg	Asn	Thr	Leu	Val	Ala	Ile	Val	Val	Gly	Val	Gly
				85					90					95	
Arg	Leu	Ile	Thr	Gly	Met	Asp	Arg	Gly	Leu	Met	Gly	Met	Cys	Val	Asn
			100					105					110		
Glu	Arg	Arg	Arg	Leu	Ile	Val	Pro	Pro	His	Leu	Gly	Tyr	Gly	Ser	Ile
		115					120					125			
Gly	Leu	Ala	Gly	Leu	Ile	Pro	Pro	Asp	Ala	Thr	Leu	Tyr	Phe	Asp	Val
	130					135					140				
Val	Leu	Leu	Asp	Val	Trp	Asn	Lys	Glu	Asp	Thr	Val	Gln	Val	Ser	Thr
145					150					155					160
Leu	Leu	Arg	Pro	Pro	His	Cys	Pro	Arg	Met	Val	Gln	Asp	Gly	Asp	Phe
				165					170					175	
Val	Arg	Tyr	His	Tyr	Asn	Gly	Thr	Leu	Leu	Asp	Gly	Thr	Ser	Phe	Asp
			180					185					190		
Thr	Ser	Tyr	Ser	Lys	Gly	Gly	Thr	Tyr	Asp	Thr	Tyr	Val	Gly	Ser	Gly
		195					200					205			
Trp	Leu	Ile	Lys	Gly	Met	Asp	Gln	Gly	Leu	Leu	Gly	Met	Cys	Pro	Gly
	210					215					220				
Gln	Arg	Arg	Lys	Ile	Ile	Ile	Pro	Pro	Phe	Leu	Ala	Tyr	Gly	Glu	Lys
225				230						235					240
Gly	Tyr	Gly	Glu	Gly	Gly	Gln	Gly	His	Lys	Gly	Lys	Phe	Arg	Arg	Arg
				245				250					255		
Gly	Lys	Asn	Gln	Ala	Ser	Thr	Tyr	Ser	Cys	Ser	Gly	Cys	Ile	Leu	His
			260					265					270		
Glu	Gly	Ile	Gln	Pro	Arg	Thr	Gln	Gly	Gly	Met	Lys	Ser	Thr	Leu	Gly
		275					280					285			
Ala	Thr	Lys	Lys	Gly	Cys	Phe	Gly	Arg	Ala	Trp	Trp	Leu	Thr	Leu	Val
	290					295					300				
Ile	Pro	Ala	Leu	Trp	Glu	Ala	Lys	Ala	Gly	Gly	Ser				
305					310					315					

```
<220>  
<221> CDS  
<222> (2)..(1324)
```

c	ttc	gat	gtg	gtt	ctg	ctg	gat	gtg	tgg	aac	aag	gaa	gac	acc	gtg	cag	49
Phe	Asp	Val	Val	Leu	Leu	Asp	Val	Trp	Asn	Lys	Glu	Asp	Thr	Val	Gln		
1				5					10					15			
gtg	agc	aca	ttg	ctg	cgc	ccg	ccc	cac	tgc	ccc	cgc	atg	gtc	cag	gac		97
Val	Ser	Thr	Leu	Leu	Arg	Pro	Pro	His	Cys	Pro	Arg	Met	Val	Gln	Asp		
			20					25					30				
ggc	gac	ttt	gtc	cgc	tac	cac	tac	aat	ggc	acc	ctg	ctg	gac	ggc	acc		145
Gly	Asp	Phe	Val	Arg	Tyr	His	Tyr	Asn	Gly	Thr	Leu	Leu	Asp	Gly	Thr		
		35					40					45					
tcc	ttc	gac	acc	agc	tac	agt	aag	ggc	ggc	act	tat	gac	acc	tac	gtc		193
Ser	Phe	Asp	Thr	Ser	Tyr	Ser	Lys	Gly	Gly	Thr	Tyr	Asp	Thr	Tyr	Val		
	50					55					60						
ggc	tct	ggt	tgg	ctg	atc	aag	ggc	atg	gac	cag	ggg	ctg	ctg	ggc	atg		241
Gly	Ser	Gly	Trp	Leu	Ile	Lys	Gly	Met	Asp	Gln	Gly	Leu	Leu	Gly	Met		
65				70					75						80		
tgt	cct	gga	gag	aga	agg	aag	att	atc	atc	cct	cca	ttc	ctg	gcc	tat		289
Cys	Pro	Gly	Glu	Arg	Arg	Lys	Ile	Ile	Ile	Pro	Pro	Phe	Leu	Ala	Tyr		
				85				90						95			
ggc	gag	aaa	ggc	tat	ggg	aca	gtg	atc	ccc	cca	cag	gcc	tgc	ctg	gtc		337
Gly	Glu	Lys	Gly	Tyr	Gly	Thr	Val	Ile	Pro	Pro	Gln	Ala	Ser	Leu	Val		
			100				105						110				
ttt	cac	gtc	ctc	ctg	att	gac	gtg	cac	aac	ccg	aag	gac	gct	gtc	cag		385
Phe	His	Val	Leu	Leu	Ile	Asp	Val	His	Asn	Pro	Lys	Asp	Ala	Val	Gln		
		115				120						125					
cta	gag	acg	ctg	gag	ctc	ccc	ccc	ggc	tgt	gtc	cgc	aga	gcc	ggg	gcc		433
Leu	Glu	Thr	Leu	Glu	Leu	Pro	Pro	Gly	Cys	Val	Arg	Arg	Ala	Gly	Ala		
	130					135					140						
ggg	gac	ttc	atg	cgc	tac	cac	tac	aat	ggc	tcc	ttg	atg	gac	ggc	acc		481
Gly	Asp	Phe	Met	Arg	Tyr	His	Tyr	Asn	Gly	Ser	Leu	Met	Asp	Gly	Thr		
145				150					155					160			
ctc	ttc	gat	tcc	agc	tac	tcc	cac	aac	cac	acc	tac	aat	acc	tat	atc		529
Leu	Phe	Asp	Ser	Ser	Tyr	Ser	His	Asn	His	Thr	Tyr	Asn	Thr	Tyr	Ile		
				165				170						175			
ggg	cag	ggt	tac	atc	atc	ccc	ggg	atg	gac	cag	ggg	ctg	cag	ggt	gcc		577
Gly	Gln	Gly	Tyr	Ile	Ile	Pro	Gly	Met	Asp	Gln	Gly	Leu	Gln	Gly	Ala		
			180				185						190				
tgc	atg	ggg	gaa	cgc	cgg	aga	att	acc	atc	ccc	ccg	cac	ctc	gcc	tat		625
Cys	Met	Gly	Glu	Arg	Arg	Arg	Ile	Thr	Ile	Pro							

ggg Gly	gag Glu	aat Asn	gga Gly	act Thr	gga Gly	gac Asp	aag Lys	atc Ile	cct Pro	ggc Gly	tct Ser	gcc Ala	gtg Val	cta Leu	atc Ile	673
210215220																
ttc Phe	aac Asn	gtc Val	cat His	gtc Val	att Ile	gac Asp	ttc Phe	cac His	aac Asn	cct Pro	gcg Ala	gat Asp	gtg Val	gtg Val	gaa Glu	721
225230235																
atc Ile	agg Arg	aca Thr	ctg Leu	tcc Ser	cgg Arg	cca Pro	tct Ser	gag Glu	acc Thr	tgc Cys	aat Asn	gag Glu	acc Thr	acc Thr	aag Lys	769
245250255																
ctt Leu	ggg Gly	gac Asp	ttt Phe	gtt Val	cga Arg	tac Tyr	cat His	tac Tyr	aac Asn	tgt Cys	tct Ser	ttg Leu	ctg Leu	gac Asp	ggc Gly	817
260265270																
acc Thr	cag Gln	ctg Leu	ttc Phe	acc Thr	tcg Ser	cat His	gac Asp	tac Tyr	ggg Gly	gcc Ala	ccc Pro	cag Gln	gag Glu	gcg Ala	act Thr	865
275280285																
ctc Leu	ggg Gly	gcc Ala	aac Asn	aag Lys	gtg Val	atc Ile	gaa Glu	ggc Gly	ctg Leu	gac Asp	acg Thr	ggc Gly	ctg Leu	cag Gln	ggc Gly	913
290295300																
atg Met	tgt Cys	gtg Val	gga Gly	gag Glu	agg Arg	cgg Arg	cag Gln	ctc Leu	atc Ile	gtg Val	ccc Pro	ccg Pro	cac His	ctg Leu	gcc Ala	961
305310315																
cac His	ggg Gly	gag Glu	agt Ser	gga Gly	gcc Ala	cgg Arg	gga Gly	gtc Val	cca Pro	ggc Gly	agt Ser	gct Ala	gtg Val	ctg Leu	ctg Leu	1009
325330335																
ttt Phe	gag Glu	gtg Val	gag Glu	ctg Leu	gtg Val	tcc Ser	cgg Arg	gag Glu	gat Asp	ggg Gly	ctg Leu	ccc Pro	aca Thr	ggc Gly	tac Tyr	1057
340345350																
ctg Leu	ttt Phe	gtg Val	tgg Trp	cac His	aag Lys	gac Asp	cct Pro	cct Pro	gcc Ala	aac Asn	ctg Leu	ttt Phe	gaa Glu	gac Asp	ata Ile	1105
355360365																
gac Asp	ctc Leu	aac Asn	aag Lys	gat Asp	ggc Gly	gag Glu	gtc Val	cct Pro	ccg Pro	gag Glu	gag Glu	ttc Phe	tcc Ser	acc Thr	ttc Phe	1153
370375380																
atc Ile	aag Lys	gct Ala	caa Gln	gtg Val	agt Ser	gag Glu	ggc Gly	aaa Lys	gga Gly	cgc Arg	ctc Leu	atg Met	cct Pro	ggg Gly	cag Gln	1201
385390395																
gac Asp	cct Pro	gag Glu	aaa Lys	acc Thr	ata Ile	gga Gly	gac Asp	atg Met	ttc Phe	cag Gln	aac Asn	cag Gln	gac Asp	cgc Arg	aac Asn	1249
405410415																
cag Gln	gac Asp	ggc Gly	aag Lys	atc Ile	aca Thr	gtc Val	gac Asp	gag Glu	ctc Leu	aag Lys	ctg Leu	aag Lys	tca Ser	gat Asp	gag Glu	1297
420425430																
gac Asp	gag Glu	gag Glu	cgg Arg	gtc Val	cac His	gag Glu	gag Glu	ctc Leu	tgaggggcag	ggagcctggc						1344
435440																
caggcctgag acacagaggc ccaactgcgag ggggacagtg gcggtggggac tgacctgctg 1404																

acagtcaccc tccctctgct gggatgaggt ccaggagcca actaaaacaa tggcagagga 1464
gacatctctg gtgttccac caccctagat gaaaatccac agcacagacc tctaccgtgt 1524
ttctcttcca tccctaaacc acttccttaa aatgtttgga ttgcaaagc caatttgggg 1584
cctgtggagc ctgggggttg atagggccat ggctgggtccc ccaccatacc tcccctccac 1644
atcactgaca cagctgagct tgttatccat ctccccaac tttctcttct tttgtacttc 1704
ttgtcatccc cactcccagc ccctattcct ctatgtgaca gctggctagg acccctctgc 1764
cttcctcccc aatcctgact ggctcctagg gaaggggaag gctcctggag ggcagcccta 1824
cctctcccat gccctttgcc ctctccctc gcctccagtg gaggctgagc tgaccctggg 1884
ctgctggagg ccagactggg ctgtagttag cttttcatcc ctaaagaagg ctttccctaa 1944
ggaaccatag aagagaggaa gaaaacaaag ggcattgtgt aggggaagctg cttgggtggg 2004
tgttagggct atgaaatctt ggatttgggg ctgaggggtg ggagggaggg cagagctctg 2064
cacactcaaa ggctaaactg gtgtcagtc ttttttcctt tgttccaaat aaaagattaa 2124
accaaaaaaa aaaaaaaaaa a 2145

<210> 4

<211> 441

<212> PRT

<213> Homo sapiens

<400> 4

Phe Asp Val Val Leu Leu Asp Val Trp Asn Lys Glu Asp Thr Val Gln
1 5 10 15

Val Ser Thr Leu Leu Arg Pro Pro His Cys Pro Arg Met Val Gln Asp
20 25 30

Gly Asp Phe Val Arg Tyr His Tyr Asn Gly Thr Leu Leu Asp Gly Thr
35 40 45

Ser Phe Asp Thr Ser Tyr Ser Lys Gly Gly Thr Tyr Asp Thr Tyr Val
50 55 60

Gly Ser Gly Trp Leu Ile Lys Gly Met Asp Gln Gly Leu Leu Gly Met
65 70 75 80

Cys Pro Gly Glu Arg Arg Lys Ile Ile Ile Pro Pro Phe Leu Ala Tyr
85 90 95

Gly Glu Lys Gly Tyr Gly Thr Val Ile Pro Pro Gln Ala Ser Leu Val
100 105 110

Phe His Val Leu Leu Ile Asp Val His Asn Pro Lys Asp Ala Val Gln
115 120 125

Leu Glu Thr Leu Glu Leu Pro Pro Gly Cys Val Arg Arg Ala Gly Ala
130 135 140

Gly Asp Phe Met Arg Tyr His Tyr Asn Gly Ser Leu Met Asp Gly Thr
145 150 155 160

```
<210> 5
<211> 3451
<212> DNA
<213> Homo sapiens
<220>
```

```

<221> CDS
<222> (130)..(1752)

<220>
<221> misc_feature
<222> (1124)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (2061)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (2111)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (2171)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (2411)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (2449)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (2451)..(2452)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (2470)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (2475)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (2477)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (2485)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (2520)
<223> n equals a, t, g, or c

```

<220>
 <221> misc_feature
 <222> (2529)
 <223> n equals a, t, g, or c

<220>
 <221> misc_feature
 <222> (2570)
 <223> n equals a, t, g, or c

<220>
 <221> misc_feature
 <222> (2581)
 <223> n equals a, t, g, or c

<220>
 <221> misc_feature
 <222> (2589)
 <223> n equals a, t, g, or c

<220>
 <221> misc_feature
 <222> (2600)
 <223> n equals a, t, g, or c

<220>
 <221> misc_feature
 <222> (2602)
 <223> n equals a, t, g, or c

<220>
 <221> misc_feature
 <222> (2604)
 <223> n equals a, t, g, or c

<220>
 <221> misc_feature
 <222> (2606)
 <223> n equals a, t, g, or c

<220>
 <221> misc_feature
 <222> (2611)
 <223> n equals a, t, g, or c

<220>
 <221> misc_feature
 <222> (2613)
 <223> n equals a, t, g, or c

<400> 5
 gcgtccgcgg ctgcagcccg ggtaggccca ggagaccggt tccacgtttg caaacgcagc 60
 cgaacgccca ggccgaccg tgccgcccga gcgcgcgcgt gcgtccgcgc cactcttctc 120
 gccgccccg atg gcg ttc cgg ggc tgg agg ccc ccg ccg cca ccg ctg ctc 171
 Met Ala Phe Arg Gly Trp Arg Pro Pro Pro Pro Leu Leu
 1 5 10

ctg ctg ctg ctc tgg gtg acc ggg cag gca gcg ccc gtg gcg ggc ctg 219
 Leu Leu Leu Leu Trp Val Thr Gly Gln Ala Ala Pro Val Ala Gly Leu

15					20					25					30					
ggc	tcc	gac	gcg	gag	ctg	cag	atc	gag	cgg	cgc	ttc	gtg	ccc	gac	gag	267				
Gly	Ser	Asp	Ala	Glu	Leu	Gln	Ile	Glu	Arg	Arg	Phe	Val	Pro	Asp	Glu					
				35					40					45						
tgc	ccg	cgc	acc	gtg	cgc	agc	ggc	gac	ttc	gtg	cgc	tac	cac	tac	gtg	315				
Cys	Pro	Arg	Thr	Val	Arg	Ser	Gly	Asp	Phe	Val	Arg	Tyr	His	Tyr	Val					
			50					55					60							
ggg	acg	ttc	ccc	gac	ggc	cag	aag	ttc	gac	tcc	agc	tat	gac	aga	gac	363				
Gly	Thr	Phe	Pro	Asp	Gly	Gln	Lys	Phe	Asp	Ser	Ser	Tyr	Asp	Arg	Asp					
		65					70					75								
tcc	act	ttc	aat	gtg	ttt	gtg	gga	aaa	gga	cag	ctg	atc	aca	ggg	atg	411				
Ser	Thr	Phe	Asn	Val	Phe	Val	Gly	Lys	Gly	Gln	Leu	Ile	Thr	Gly	Met					
	80					85					90									
gac	cag	gct	ctt	gtt	ggg	atg	tgc	gta	aac	gag	aga	cgt	ttc	gtg	aag	459				
Asp	Gln	Ala	Leu	Val	Gly	Met	Cys	Val	Asn	Glu	Arg	Arg	Phe	Val	Lys					
95					100				105						110					
att	ccc	cca	aag	ctt	gcc	tac	gga	aat	gaa	aga	gtt	tct	ggt	gtg	atc	507				
Ile	Pro	Pro	Lys	Leu	Ala	Tyr	Gly	Asn	Glu	Arg	Val	Ser	Gly	Val	Ile					
				115				120						125						
ccc	ccc	aat	tca	gtg	ctt	cat	ttt	gat	gta	ctt	ctg	atg	gat	att	tgg	555				
Pro	Pro	Asn	Ser	Val	Leu	His	Phe	Asp	Val	Leu	Leu	Met	Asp	Ile	Trp					
			130					135					140							
aat	tct	gaa	gac	cag	gtt	cag	att	cac	acc	tat	ttc	aag	ccc	ccg	agt	603				
Asn	Ser	Glu	Asp	Gln	Val	Gln	Ile	His	Thr	Tyr	Phe	Lys	Pro	Pro	Ser					
		145					150					155								
tgc	cct	cgg	acc	atc	cag	gtg	tct	gat	ttt	gtg	agg	tac	cac	tac	aac	651				
Cys	Pro	Arg	Thr	Ile	Gln	Val	Ser	Asp	Phe	Val	Arg	Tyr	His	Tyr	Asn					
	160					165					170									
ggg	acg	ttc	ctg	gac	gga	act	ctg	ttt	gat	tgc	agt	cac	aat	cgc	atg	699				
Gly	Thr	Phe	Leu	Asp	Gly	Thr	Leu	Phe	Asp	Ser	Ser	His	Asn	Arg	Met					
175					180				185						190					
aaa	aca	tat	gac	acg	tat	gtg	gga	att	ggc	tgg	ctg	att	cct	gga	atg	747				
Lys	Thr	Tyr	Asp	Thr	Tyr	Val	Gly	Ile	Gly	Trp	Leu	Ile	Pro	Gly	Met					
				195					200					205						
gat	aaa	ggg	ctg	ctg	ggg	atg	tgt	gtg	ggt	gag	aag	cgc	atc	atc	acc	795				
Asp	Lys	Gly	Leu	Leu	Gly	Met	Cys	Val	Gly	Glu	Lys	Arg	Ile	Ile	Thr					
			210					215					220							
att	cct	cct	ttt	ctg	gcc	tat	gga	gag	gat	gga	gat	ggg	aaa	gac</						

tgt gag cgg ata agt caa agt ggg gac ttt ctc acg tat cat tac aat	987
Cys Glu Arg Ile Ser Gln Ser Gly Asp Phe Leu Thr Tyr His Tyr Asn	
275 280 285	
ggc acg ctt ctg gat ggc acc ctc ttt gat tcc agc tac tct cgg aac	1035
Gly Thr Leu Leu Asp Gly Thr Leu Phe Asp Ser Ser Tyr Ser Arg Asn	
290 295 300	
cgc acg ttt gac acg tac att ggg cag ggc tac gtg att cct ggg atg	1083
Arg Thr Phe Asp Thr Tyr Ile Gly Gln Gly Tyr Val Ile Pro Gly Met	
305 310 315	
gat gaa ggt cta ctt ggt gtt tgc att gga gaa aag cga ang att gtg	1131
Asp Glu Gly Leu Leu Gly Val Cys Ile Gly Glu Lys Arg Xaa Ile Val	
320 325 330	
gtc ccg cct cac ctg ggg tat gga gag gaa gga aga ggg aat atc ccc	1179
Val Pro Pro His Leu Gly Tyr Gly Glu Glu Gly Arg Gly Asn Ile Pro	
335 340 345 350	
ggc tcg gct gtg ctg gtg ttt gac atc cat gtg atc gac ttc cac aac	1227
Gly Ser Ala Val Leu Val Phe Asp Ile His Val Ile Asp Phe His Asn	
355 360 365	
cct tcg gac tcc atc agc atc acc tcc cac tac aaa ccc cct gac tgc	1275
Pro Ser Asp Ser Ile Ser Ile Thr Ser His Tyr Lys Pro Pro Asp Cys	
370 375 380	
tca gtg ctg agt aag aag gga gat tac ctc aaa tat cac tac aat gcc	1323
Ser Val Leu Ser Lys Lys Gly Asp Tyr Leu Lys Tyr His Tyr Asn Ala	
385 390 395	
tca ctt ctg gat ggg acc ctg ctg gac tcc acg tgg aat tta ggc aaa	1371
Ser Leu Leu Asp Gly Thr Leu Leu Asp Ser Thr Trp Asn Leu Gly Lys	
400 405 410	
act tac aat att gtt ctg gga tct ggg caa gtt gtg ttg ggg atg gac	1419
Thr Tyr Asn Ile Val Leu Gly Ser Gly Gln Val Val Leu Gly Met Asp	
415 420 425 430	
atg ggt ctc aga gag atg tgc gtt ggc gag aaa cgg aca gtg atc att	1467
Met Gly Leu Arg Glu Met Cys Val Gly Glu Lys Arg Thr Val Ile Ile	
435 440 445	
ccg cct cac ctg ggc tat ggg gaa gct ggc gtg gat gga gaa gtg ccc	1515
Pro Pro His Leu Gly Tyr Gly Glu Ala Gly Val Asp Gly Glu Val Pro	
450 455 460	
ggc agt gcc gta tta gtg ttt gac att gag cts ctg gag ctg gtg gct	1563
Gly Ser Ala Val Leu Val Phe Asp Ile Glu Xaa Leu Glu Leu Val Ala	
465 470 475	
ggc ctt cct gag ggg tac atg ttc ata tgg aat ggt gag gtg tca ccc	1611
Gly Leu Pro Glu Gly Tyr Met Phe Ile Trp Asn Gly Glu Val Ser Pro	
480 485 490	
aac ctc ttt gaa gaa att gac aag gat ggc aac gga gaa gtc ctc ctg	1659
Asn Leu Phe Glu Glu Ile Asp Lys Asp Gly Asn Gly Glu Val Leu Leu	
495 500 505 510	
gaa gag ttc tca gag tac att cac gcc cag gtg gca tct ggc aaa ggg	1707
Glu Glu Phe Ser Glu Tyr Ile His Ala Gln Val Ala Ser Gly Lys Gly	

515								520				525					
aaa ctc gct cct ggc ttt gat gct gag ctg att gtg aag aat atg																	1752
Lys Leu Ala Pro Gly Phe Asp Ala Glu Leu Ile Val Lys Asn Met																	
530								535				540					
ttcaccaacc aggaccggaa tggagatggg aaggtcacag ccgaggaatt taaactcara																	1812
gaccaggaag ccaaacacga tgtaactcta aacctggcat gaaccagatg gtgccagggg																	1872
agtacgtgac accaagccac ctgtgtggma agacgtgcar ttgarggtgc aaggggtctc																	1932
tcagaagttt gcatccatta gccagtagta ggtgggggtca catagtacct ggtgtacaca																	1992
tcgggggtggg ttgatatatg gggtgagaag tttgggctga tcgccagtga tagtaaaaa																	2052
aatctgtgnc agagggcctt agcatgggat gtgtccagta ttcgaccac gcgtccgcnt																	2112
gccaaecatg atttgtgagc cttctgggaa attttgttat taaaggaata tatagtgtnc																	2172
agacggaagt tataatcatc ttggaggaac cataagaaaa ggtgtccagg gtatctatat																	2232
aaagaggggt aaattttttt ttaacttgct ggttaaaaaca ttttagaaat attctagaga																	2292
tgggcaggag agtcaaaggg cttgcttgcc ccagcagagt tcccagcaga cagccatgga																	2352
tattcccagc agcctgtgca aattctgatg atggcccccac ccccgcacac ggcacacgnc																	2412
acatcawgct tttccagctc atcacacccc gcccantnn gggcctacca ttaatagngt																	2472
atnanttga ggntaaaaga gccttttgga cagaaaactg ggccaggnaa aggcatntca																	2532
gaccacaaat agagaatttg attcgtcatt tgccacanaa gtcatctgnt tagcttntcc																	2592
tttccttnan ananatttna ntttttctg gaggcagagt ctccctttgt cgccaggctg																	2652
gagtgcagtg gtgccatctc ggctcactgc agcactgtct cggctcactg cagcktcogc																	2712
ctcccgattt caagcgattc tcctgtctca gcctcctgag tagctgggas tacagggtgtg																	2772
caccaccacg ccoggctaatt ttttgtattt ttagtagaga cgggggtttca ctgtgttggc																	2832
caggatggtc tcaatctcga cttcgtgatc cgcccacctt ggccctccaa agtgttggga																	2892
ttacaggcgt gactcaccat gccagccac ttagtttttt cttattccca cttttctatc																	2952
ccatagaaca ttttttttat cttccctgaa ccataatwgat gagataaata gggctggggg																	3012
mtgggccccg ctgggtcactc aacagagtat ttcccttggc cgagatggaa gttttgtccc																	3072
aatagatgag ctgctgagta tcaacaagggt gacatttttc tgctgcccatt ttgtgtcctg																	3132
gagacgggtg taccctgaag gcagaggcca gtgcgcgaag acagcaatga cagtccaact																	3192
gccgacctga ttctgtcatc atggaataac cacatggcta ctttctatcc tctgttccca																	3252
aatggtgggtg gcacttatcc tgaagtcgtc aatgaacttc ctttgaaact actttatttt																	3312
actaatttaa actattttgt actgatgtag ccctgaggta gttcatgaaa atgctgtgca																	3372
ctcattccat ggaataaatg ttggaaagct catcttttct gatataaaat gttgaatgat																	3432

3451

<400>	6															
Met	Ala	Phe	Arg	Gly	Trp	Arg	Pro	Pro	Pro	Pro	Pro	Leu	Leu	Leu	Leu	
1				5					10					15		
Leu	Leu	Trp	Val	Thr	Gly	Gln	Ala	Ala	Pro	Val	Ala	Gly	Leu	Gly	Ser	
			20					25					30			
Asp	Ala	Glu	Leu	Gln	Ile	Glu	Arg	Arg	Phe	Val	Pro	Asp	Glu	Cys	Pro	
		35					40					45				
Arg	Thr	Val	Arg	Ser	Gly	Asp	Phe	Val	Arg	Tyr	His	Tyr	Val	Gly	Thr	
	50					55					60					
Phe	Pro	Asp	Gly	Gln	Lys	Phe	Asp	Ser	Ser	Tyr	Asp	Arg	Asp	Ser	Thr	
65					70					75					80	
Phe	Asn	Val	Phe	Val	Gly	Lys	Gly	Gln	Leu	Ile	Thr	Gly	Met	Asp	Gln	
				85					90					95		
Ala	Leu	Val	Gly	Met	Cys	Val	Asn	Glu	Arg	Arg	Phe	Val	Lys	Ile	Pro	
			100					105					110			
Pro	Lys	Leu	Ala	Tyr	Gly	Asn	Glu	Arg	Val	Ser	Gly	Val	Ile	Pro	Pro	
		115					120					125				
Asn	Ser	Val	Leu	His	Phe	Asp	Val	Leu	Leu	Met	Asp	Ile	Trp	Asn	Ser	
	130					135					140					
Glu	Asp	Gln	Val	Gln	Ile	His	Thr	Tyr	Phe	Lys	Pro	Pro	Ser	Cys	Pro	
145					150					155					160	
Arg	Thr	Ile	Gln	Val	Ser	Asp	Phe	Val	Arg	Tyr	His	Tyr	Asn	Gly	Thr	
				165					170					175		
Phe	Leu	Asp	Gly	Thr	Leu	Phe	Asp	Ser	Ser	His	Asn	Arg	Met	Lys	Thr	
			180					185					190			
Tyr	Asp	Thr	Tyr	Val	Gly	Ile	Gly	Trp	Leu	Ile	Pro	Gly	Met	Asp	Lys	
		195					200					205				
Gly	Leu	Leu	Gly	Met	Cys	Val	Gly	Glu	Lys	Arg	Ile	Ile	Thr	Ile	Pro	
	210					215					220					
Pro	Phe	Leu	Ala	Tyr	Gly	Glu	Asp	Gly	Asp	Gly	Lys	Asp	Ile	Pro	Gly	
225					230					235					240	
Gln	Ala	Ser	Leu	Val	Phe	Asp	Val	Ala	Leu	Leu	Asp	Leu	His	Asn	Pro	
				245					250					255		
Lys	Asp	Ser	Ile	Ser	Ile	Glu	Asn	Lys	Val	Val	Pro	Glu	Asn	Cys	Glu	
			260					265					270			
Arg	Ile	Ser	Gln	Ser	Gly	Asp	Phe	Leu	Thr	Tyr	His	Tyr	Asn	Gly	Thr	
		275					280					285				

```
<210> 7
<211> 1251
<212> DNA
<213> Homo sapiens
```

<400> 7
ac acg tat ggg gaa att ggc tgg ctg att cct gga atg gat aaa ggg
Thr Tyr Gly Glu Ile Gly Trp Leu Ile Pro Gly Met Asp Lys Gly

1	5	10	15	
ctg ctg ggg atg tgt gtg ggt gag aag cgc atc atc acc att cct cct				95
Leu Leu Gly Met Cys Val Gly Glu Lys Arg Ile Ile Thr Ile Pro Pro	20	25	30	
ttt ctg gcc tat gga gag gat gga gat ggg aaa gac att ccc ggt cag				143
Phe Leu Ala Tyr Gly Glu Asp Gly Asp Gly Lys Asp Ile Pro Gly Gln	35	40	45	
gca tct ctg gtg ttt gat gtt gca tta ttg gac ctc cat aac ccc aag				191
Ala Ser Leu Val Phe Asp Val Ala Leu Leu Asp Leu His Asn Pro Lys	50	55	60	
gac agc att tcc att gag aac aag gta gta cct gaa aac tgt gag cgg				239
Asp Ser Ile Ser Ile Glu Asn Lys Val Val Pro Glu Asn Cys Glu Arg	65	70	75	
ata agt caa agt ggg gac ttt ctc agg tat cat tac aat ggc acg ctt				287
Ile Ser Gln Ser Gly Asp Phe Leu Arg Tyr His Tyr Asn Gly Thr Leu	80	85	90	95
ctg gat ggc acc ctc ttt gat tcc agc tac tct cgg aac cgc acg ttt				335
Leu Asp Gly Thr Leu Phe Asp Ser Ser Tyr Ser Arg Asn Arg Thr Phe	100	105	110	
gac acg tac att ggg cag ggc tac gtg att cct ggg atg gat gaa ggt				383
Asp Thr Tyr Ile Gly Gln Gly Tyr Val Ile Pro Gly Met Asp Glu Gly	115	120	125	
cta ctt ggt gtt tgc att gga gaa aag cga agg att gtg gtc ccg cct				431
Leu Leu Gly Val Cys Ile Gly Glu Lys Arg Arg Ile Val Val Pro Pro	130	135	140	
cac ctg ggg tat gga gag gaa gga aga ggg aat atc ccc ggc tcg gct				479
His Leu Gly Tyr Gly Glu Glu Gly Arg Gly Asn Ile Pro Gly Ser Ala	145	150	155	
gtg ctg gtg ttt gac atc cat gtg atc gac ttc cac aac cct tcg gac				527
Val Leu Val Phe Asp Ile His Val Ile Asp Phe His Asn Pro Ser Asp	160	165	170	175
tcc atc agc atc acc tcc cac tac aaa ccc cct gac tgc tca gtg ctg				575
Ser Ile Ser Ile Thr Ser His Tyr Lys Pro Pro Asp Cys Ser Val Leu	180	185	190	
agt aag aag gga gat tac ctc aaa tat cac tac aat gcc tca ctt ctg				623
Ser Lys Lys Gly Asp Tyr Leu Lys Tyr His Tyr Asn Ala Ser Leu Leu	195	200	205	
gat ggg acc ctg ctg gac tcc acg tgg aat tta ggc aaa act tac aat				671
Asp Gly Thr Leu Leu Asp Ser Thr Trp Asn Leu Gly Lys Thr Tyr Asn	210	215	220	
att gtt ctg gga tct ggg caa gtt gtg ttg ggg atg gac atg ggt ctc				719
Ile Val Leu Gly Ser Gly Gln Val Val Leu Gly Met Asp Met Gly Leu	225	230	235	
aga gag atg tgc gtt ggc gag aaa cgg aca gtg atc att ccg cct cac				767
Arg Glu Met Cys Val Gly Glu Lys Arg Thr Val Ile Ile Pro Pro His	240	245	250	255

```

<400> 8
Thr Tyr Gly Glu Ile Gly Trp Leu Ile Pro Gly Met Asp Lys Gly Leu
  1           5           10           15
Leu Gly Met Cys Val Gly Glu Lys Arg Ile Ile Thr Ile Pro Pro Phe
           20           25           30
Leu Ala Tyr Gly Glu Asp Gly Asp Gly Lys Asp Ile Pro Gly Gln Ala
           35           40           45
Ser Leu Val Phe Asp Val Ala Leu Leu Asp Leu His Asn Pro Lys Asp
           50           55           60
Ser Ile Ser Ile Glu Asn Lys Val Val Pro Glu Asn Cys Glu Arg Ile
           65           70           75           80
Ser Gln Ser Gly Asp Phe Leu Arg Tyr His Tyr Asn Gly Thr Leu Leu
           85           90           95

```

[illegible]

3/10

FIGURE 2A

HL1AP03

10 30 50
CTTCGATGTGGTTCTGCTGGATGTGTGGAACAAGGAAGACACCGTGCAGGTGAGCACATT
F D V V L L D V W N K E D T V Q V S T L
70 90 110
GCTGCGCCCGCCCCACTGCCCCCGCATGGTCCAGGACGGCGACTTTGTCCGCTACCACTA
L R P P H C P R M V Q D G D F V R Y H Y
130 150 170
CAATGGCACCCCTGCTGGACGGCACCTCCTTCGACACCAGCTACAGTAAGGGCGGCACTTA
N G T L L D G T S F D T S Y S K G G T Y
190 210 230
TGACACCTACGTCGGCTCTGGTTGGCTGATCAAGGGCATGGACCAGGGGCTGCTGGGCAT
D T Y V G S G W L I K G M D Q G L L G M
250 270 290
GTGTCCTGGAGAGAGAAGGAAGATTATCATCCCTCCATTCCCTGGCCTATGGCGAGAAAGG
C P G E R R K I I I P P F L A Y G E K G
310 330 350
CTATGGGACAGTGATCCCCCACAGGCCTCGCTGGTCTTTACGTCCTCCTGATTGACGT
Y G T V I P P Q A S L V F H V L L I D V
370 390 410
GCACAACCCGAAGGACGCTGTCCAGCTAGAGACGCTGGAGCTCCCCCGGCTGTGTCCG
H N P K D A V Q L E T L E L P P G C V R
430 450 470
CAGAGCCGGGGCGGGGACTTCATGCGCTACCACTACAATGGCTCCTTGATGGACGGCAC
R A G A G D F M R Y H Y N G S L M D G T
490 510 530
CCTCTTCGATTCCAGCTACTCCCACAACCACCTACAATACCTATATCGGGCAGGGTTA
L F P S S Y S H N H T Y N T Y I G Q G Y
550 570 590
CATCATCCCCGGGATGGACCAGGGGCTGCAGGGTGCCTGCATGGGGGAACGCCGGAGAAT
I I P G M D Q G L Q G A C M G E R R R I
610 630 650
TACCATCCCCCGCACCTCGCCTATGGGGAGAATGGAAGTGGAGACAAGATCCCTGGCTC
T I P P H L A Y G E N G T G D K I P G S
670 690 710
TGCCGTGCTAATCTCAACGTCCATGTCATTGACTTCCACAACCCTGCGGATGTGGTGA
A V L I F N V H V I D F H N P A D V V E
730 750 770
AATCAGGACACTGTCCCGGCATCTGAGACCTGCAATGAGACCACCAAGCTTGGGGACTT
I R T L S R P S E T C N E T T K L G D F

09225502 " 010699

FIGURE 2B

790 810 830
TGTTTCGATACCATTAACAAGTCTCTTTGCTGGACGGCACCCAGCTGTTTCACCTCGCATGA
V R Y H Y N C S L L D G T Q L F T S H D
850 870 890
CTACGGGGCCCCCAGGAGCGACTCTCGGGGCCAACAAGGTGATCGAAGGCCTGGACAC
Y G A P Q E A T L G A N K V I E G L D T
910 930 950
GGGCCTGCAGGGCATGTGTGTGGGAGAGAGGCGGCAGCTCATCGTGCCCCCGCACCTGGC
G L Q G M C V G E R R Q L I V P P H L A
970 990 1010
CCACGGGSAGAGTGGAGCCCCGGGAGTCCCAGGCAGTGCTGTGCTGCTGTTTGAGGTGGA
H G E S G A R G V P G S A V L L F E V E
1030 1050 1070
GCTGGTGTCCCGGGAGGATGGGCTGCCCCACAGGCTACCTGTTTGTGTGGCACAAGGACCC
L V S R E D G L P T G Y L F V W H K D P
1090 1110 1130
TCCTGCCAACCTGTTTGAAGACATAGACCTCAACAAGGATGGCGAGGTCCCTCCGGAGGA
P A N L F E D I D L N K D G E V P P E E
1150 1170 1190
GTTCTCCACCTTCATCAAGGCTCAAGTGAGTGAGGGCAAAGGACGCCTCATGCCTGGGCA
F S T F I K A Q V S E G K G R L M P G Q
1210 1230 1250
GGACCCTGAGAAAACCATAGGAGACATGTTCCAGAACCAGGACCGCAACCAGGACGGCAA
D P E K T I G D M F Q N Q D R N Q D G K
1270 1290 1310
GATCACAGTCGACGAGCTCAAGCTGAAGTCAGATGAGGACGAGGAGCGGGTCCACGAGGA
I T V D E L K L K S D E D E E R V H E E
1330 1350 1370
GCTCTGAGGGGCAGGGAGCCTGGCCAGGCCTGAGACACAGAGGCCCACTGCGAGGGGGAC
L *
1390 1410 1430
AGTGGCGGTGGGACTGACCTGCTGACAGTCACCCTCCCTCTGCTGGGATGAGGTCCAGGA
1450 1470 1490
GCCAACTAAAACAATGGCAGAGGAGACATCTCTGGTGTTCACCACCCCTAGATGAAAAT
1510 1530 1550
CCACAGCACAGACCTCTACCGTGTCTCTCTCCATCCCTAAACCACTTCCTTAAAATGTT
1570 1590 1610
TGGATTGCAAAGCCAATTTGGGGCCTGTGGAGCCTGGGGTTGGATAGGGCCATGGCTGG
1630 1650 1670
TCCCCCACCATACCTCCCCTCCACATCACTGACACAGCTGAGCTTGTATCCATCTCCCC
1690 1710 1730
AAACTTTCTCTTTCTTTGTACTTCTTGTCTATCCCCACTCCCAGCCCCCTATTCTCTATGT

09225502 010699

FIGURE 2C

1750 1770 1790
GACAGCTGGCTAGGACCCCTCTGCCTTCCTCCCCAATCCTGACTGGCTCCTAGGGAAGGG
1810 1830 1850
GAAGGCTCCTGGAGGGCAGCCCTACCTCTCCCATGCCCTTTGCCCTCCTCCCTCGCCTCC
1870 1890 1910
AGTGGAGGCTGAGCTGACCCCTGGGCTGCTGGAGGCCAGACTGGGCTGTAGTTAGCTTTTC
1930 1950 1970
ATCCCTAAGAAGGCTTTCCCTAAGGAACCATAGAAGAGAGGAAGAAAACAAAGGGCATG
1990 2010 2030
TGTGAGGGAAGCTGCTTGGGTGGGTGTTAGGGCTATGAAATCCTTGGATTTGGGGCTGAGG
2050 2070 2090
GGTGGGAGGGAGGGCAGAGCTCTGCACACTCAAAGGCTAAACTGGTGTCAGTCCTTTTTT
2110 2130
CCTTTGTTCCAAATAAAAGATTAAACCAAAAAAAAAAAAAAAAAAAAA

0922502.010699

6/10

FIGURE 3A

HSYBM46

CGGTCCGCGGCTGCAGCCCGGGTAGGGCCAGGAGACCCGGTCCACGTTTGCAAACGCAGC
CGAACGCCCAGGCCGACCCGTGCCGCCCAGCGCCGCGCTGCGTCCGCGCCACTCTTCTC

M A F R G W R P P P P P L L L L L
GCCGCCCCGATGGCGTTCGGGGCTGGAGGCCCCGCCGCCACCGCTGCTCCTGCTGCTG

L W V T G Q A A P V A G L G S D A E L Q
CTCTGGGTGACCGGCAGGCAGCGCCCGTGGCGGGCCTGGGCTCCGACGCGGAGCTGCAG

I E R R F V P D E C P R T V R S G D F V
ATCGAGCGGCGCTTCGTGCCCGACGAGTGCCCGCGCACCGTGCGCAGCGGCGACTTCGTG

R Y H Y V G T F P D G Q K F D S S Y D R
CGCTACCACTACGTGGGGACGTTCCTCCGACGGCCAGAAGTTCGACTCCAGCTATGACAGA

D S T F N V F V G K G Q L I T G M D Q A
GACTCCACTTTCAATGTGTTTGTGGGAAAAGGACAGCTGATCACAGGGATGGACCAGGCT

L V G M C V N E R R F V K I P P K L A Y
CTTGTTGGGATGTGCGTAAACGAGAGACGTTTCGTGAAGATTCCCCCAAAGCTTGCCTAC

G N E R V S G V I P P N S V L H F D V L
GGAAATGAAAGAGTTTCTGGTGTGATCCCCCCCCAATTCAGTGCTTCATTTTGATGTACTT

L M D I W N S E D Q V Q I H T Y F K P P
CTGATGGATATTTGGAATTCTGAAGACCAGGTTTCAGATTCACACCTATTTCAAGCCCCCG

S C P R T I Q V S D F V R Y H Y N G T F
AGTTGCCCTCGGACCATCCAGGTGTCTGATTTTGTGAGGTACCACTACAACGGGACGTTTC

L D G T L F D S S H N R M K T Y D T Y V
CTGGACGGAACCTCTGTTTGATTTCGAGTCACAATCGCATGAAAACATATGACACGTATGTG

G I G W L I P G M D K G L L G M C V G E
GGAATTGGCTGGCTGATTCTGGAATGGATAAAGGGCTGCTGGGGATGTGTGTGGGTGAG

K R I I T I P P F L A Y G E D G D G K D
AAGCGCATCATCACCATTCTCCTTTTCTGGCCTATGGAGAGGATGGAGATGGGAAAGAC

I P G Q A S L V F D V A L L D L H N P K
ATTCCCGGTGAGCATCTCTGGTGTGTTGATGTTGCATTATTGGACCTCCATAACCCCAAG

D S I S I E N K V V P E N C E R I S Q S
GACAGCATTTCCATTGAGAACAAGGTAGTACCTGAAAACGTGAGCGGATAAGTCAAAGT

G D F L T Y H Y N G T L L D G T L F D S
GGGGACTTTCTCACGTATCATTACAATGGCACGCTTCTGGATGGCACCTCTTTGATTCC

0922502 010699

FIGURE 3B

S Y S R N R T F D T Y I G Q G Y V I P G
 AGCTACTCTCGGAACCGCACGTTTGACACGTACATTGGGCAGGGCTACGTGATTCCTGGG

 M D E G L L G V C I G E K R X I V V P P
 ATGGATGAAGGTCTACTTGGTGT'TTGCATTGGAGAAAAGCGAANGATTGTGGTCCCGCCT

 H L G Y G E E G R G N I P G S A V L V F
 CACCTGGGGTATGGAGAGGAAGGAAGAGGGAATATCCCCGGCTCGGCTGTGCTGGTGT'TT

 D I H V I D F H N P S D S I S I T S H Y
 GACATCCATGTGATCGACTTCCACAACCCTTCGGACTCCATCAGCATCACCTCCCCTACTAC

 K P P D C S V L S K K G D Y L K Y H Y N
 AAACCCCTGACTGCTCAGTGC'TGAGTAAGAAGGGAGATTACCTCAAATATCACTACAAT

 A S L L D G T L L D S T W N L G K T Y N
 GCCTCACTTCTGGATGGGACCTGCTGGACTCCACGTGGAATTTAGGCAAACTTACAAT

 I V L G S G Q V V L G M D M G L R E M C
 ATTGTTCTGGGATCTGGGCAAGTTGTGT'TGGGGATGGACATGGGTCTCAGAGAGATGTGC

 V G E K R T V I I P P H L G Y G E A G V
 GTTGGCGAGAAACGGACAGTGATCAT'TCCGCCTCACCTGGGCTATGGGGAAGCTGGCGTG

 D G E V P G S A V L V F D I E L L E L V
 GATGGAGAAGTGCCCGGCAGTGCCGTAT'TAGTGT'TTGACATTGAGCTSCTGGAGCTGGTG

 A G L P E G Y M F I W N G E V S P N L F
 GCTGGCCTTCTCTGAGGGGTACATGTTCATATGGAATGGTGAGGTGTCACCCAACCTCTTT

 E E I D K D G N G E V L L E E F S E Y I
 GAAGAAATTGACAAGGATGGCAACGGAGAAGTCTCTGGAAGAGTTCTCAGAGTACATT

 H A Q V A S G K G K L A P G F D A E L I
 CACGCCCAGGTGGCATCTGGCAAAGGGAAACTCGCTCCTGGCTTTGATGCTGAGCTGATT

 V K N M F T N Q D R N G D G K V T A E E
 GTGAAGAATATGTTTACCAACCAGGACCGGAATGGAGATGGGAAGGTCACAGCCGAGGAA

 F K L X D Q E A K H D V T L N L A *
 TTTAAACTCARAGACCAGGAAGCCAAACACGATGTAACCTCTAAACCTGGCATGAACCAGA

 TGGTGCCAGGGGAGTACGTGACACCAAGCCACCTGTGTGGMAAGACGTGCARTTGARGGT
 GCAAGGGGTCTCTCAGAAGTTTGATCCATTAGCCAGTAGTAGGTGGGGTCACATAGTAC
 CTGGTGTACACATCGGGGTGGGT'TGATATATGGGGTGAGAAGTTTGGGCTGATCGCCAGT
 GATAGTAAACAAAATCTGTGNCAGAGGGCCTTAGCATGGGATGTGTCCAGTATTCGACCC
 ACGCGTCCGCTNGCCAACCATGATTTGTGAGCCTTCTGGGAAATTTGTATTAAGGAA
 TATATAGTGTNCAGACGGAAGTTATAATCATCTTGAGGAACCATAAGAAAAGGTGTCCA
 GGGTATCTATATAAGAGGGTTAAATTTT'TTAACTTGCTGGTTAAAACATTTTAGAA
 ATATTCTAGAGATGGGCAGGAGAGTCAAAGGGCTTGCTTGCCCCAGCAGAGTCCCAGCA
 GACAGCCATGGATAT'TCCAGCAGCCTGTGCAAATCTGATGATGGCCCCACCCCGCAC
 ACGGCACACGNCACATCAWGCTTTTCCAGCTCATCACACCCCGCCCCANTNNGGCCTAC
 CATTAATAGNGTATNANTTGGAGGNTAAAAGAGCCTTTTGGACAGAAAACCTGGGCCAGGN
 AAAGGCATNTCAGACCACAAATAGAGAATTTGATTGTCATTTGCCACANAAGTCATCTG

09225502.010699

FIGURE 3C

NTTAGCTTNTCCTTTCCTTNANANANATTTNANTTTTTCTGGAGGCAGAGTCTCCCTTT
GTCGCCAGGCTGGAGTGCAGTGGTGCCATCTCGGCTCACTGCAGCACTGTCTCGGCTCAC
TGCAGCKTCCGCCTCCCGTATTCAAGCGATTCTCCTGTCTCAGCCTCCTGAGTAGCTGGG
ASTACAGGTGTGCACCACCACGCCCCGGCTAATTTTTGTATTTTATAGTAGAGACGGGGTTT
CACTGTGTTGGCCAGGATGGTCTCAATCTCGACTTCGTGATCCGCCCACCTTGGCCTCCC
AAAGTGTGGGATTACAGGCGTGACTCACCATGCCCAGCCACTTAGTTTTTTCTTATTCC
CACCTTCTATCCCATAGAACATCTTTTTTATCTTCCCTGAACCATATWGATGAGATAAA
TAGGGCTGGGGMTGGGCCCCGCTGGTCACTCAACAGAGTATTTCCCTTGGCCGAGATGG
AAGTTTTGTCCCAATAGATGAGCTGCTGAGTATCAACAAGGTGACATTTTTCTGCTGCCC
ATTTGTGTCTGGAGACGGTGGTACCTGAAGGCAGAGGCCAGTGCCGCAAGACAGCAAT
GACAGTCCACCTGCCGACCTGATTCTGTCATCATGGAATAACCACATGGCTACCTTCTAT
CCTCTGTTCCCAAATGGTGGTGGCACTTATCCTGAAGTCGTCAATGACTTCCCTTTGAAA
CTACTTTATTTTACTAATTTAAACTATTTTGTACTGATGTAGCCCTGAGGTAGTTCATGA
AAATGCTGTGCACTCATTCATGGAATAAATGTTGGAAAGCTCATCTTTCTGATATAAA
ATGTTGAATGATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

0925502 "010699

9/10

FIGURE 4A

HFKBC47

```

      10              30              50
ACACGTATGGCGAAATTGGCTGGCTGATTCTGGAATGGATAAAGGGCTGCTGGGGATGT
  T Y G E I G W L I P G M D K G L L G M C
      70              90             110
GTGTGGGTGAGAAGCGCATCATCACCATTCTCTCTTTCTGGCCTATGGAGAGGATGGAG
  V G E K R I I T I P P F L A Y G E D G D
      130             150             170
ATGGGAAAGACATTCCCGGTCAGGCATCTCTGGTGTGTTGATGTTGCATTATTGGACCTCC
  G K D I P G Q A S L V F D V A L L D L H
      190             210             230
ATAACCCCAAGGACAGCATTTCCATTGAGAACAAGGTAGTACCTGAAAACGTGTGAGCGGA
  N P K D S I S I E N K V V P E N C E R I
      250             270             290
TAAGTCAAAGTGGGGACTTTCTCAGGTATCATTACAATGGCACGCTTCTGGATGGCACCC
  S Q S G D F L R Y H Y N G T L L D G T L
      310             330             350
TCTTTGATTCAGCTACTCTCGGAACCGCACGTTTGACACGTACATTGGGCAGGGCTACG
  F D S S Y S R N R T F D T Y I G Q G Y V
      370             390             410
TGATTCTCTGGGATGGATGAAGGTCTACTTGGTGTGTTGCATTGGAGAAAAGCGAAGGATTG
  I P G M D E G L L G V C I G E K R R I V
      430             450             470
TGGTCCCGCCTCACCTGGGGTATGGAGAGGAAGGAAGAGGGAATATCCCCGGCTCGGCTG
  V P P H L G Y G E E G R G N I P G S A V
      490             510             530
TGCTGGTGTGTTGACATCCATGTGATCGACTTCCACAACCCTTCGGACTCCATCAGCATCA
  L V F D I H V I D F H N P S D S I S I T
      550             570             590
CCTCCCCTACAAACCCCTGACTGCTCAGTCTGAGTAAGAAGGGAGATTACCTCAAAT
  S H Y K P P D C S V L S K K G D Y L K Y
      610             630             650
ATCACTACAATGCCTCACTTCTGGATGGGACCCTGCTGGACTCCACGTGGAATTTAGGCA
  H Y N A S L L D G T L L D S T W N L G K
      670             690             710
AAACTTACAAATATTGTTCTGGGATCTGGGCAAGTTGTGTTGGGGATGGACATGGGTCTCA
  T Y N I V L G S G Q V V L G M D M G L R
      730             750             770
GAGAGATGACCGTTGGCGAGAAACGGACAGTGATCATTCCGCCTCACCTGGGCTATGGGG
  E M V G E K R T V I I P P H L G Y G E

```

09225502 "010699

10/10

FIGURE 4B

```

      790                      810                      830
AAGCTGGCGTGGATGGAGAAAGTGCCCGGCAGTGCCGTATTAGTGTTTGACATTGAGCTGC
  A  G  V  D  G  E  V  P  G  S  A  V  L  V  F  D  I  E  L  L
      850                      870                      890
TGGAGCTGGTGGCTGGCCTTCCTGAGGGGTACATGTTTCATATGGAATGGTGAGGTGTCAC
  E  L  V  A  G  L  P  E  G  Y  M  F  I  W  N  G  E  V  S  P
      910                      930                      950
CCAACCTCTTTGAAGAAATCAACAAGGTGACATTTTTCTGCTGCCCATTGTGTCCTGGA
  N  L  E  E  I  N  K  V  T  F  F  C  C  P  F  V  S  W  R
      970                      990                      1010
GACGGTGGTACCCTGAAGGCAGAGGCCAGCTGCCGCAAGACAGCAATGACAGTCCACCTG
  R  W  Y  P  E  G  R  G  Q  L  P  Q  D  S  N  D  S  P  P  A
      1030                      1050                      1070
CCGACCTGATTTCCTGCATCATGGAATAACCACATGGCTACCTTCTATCCTCTGTTCCCAA
  D  L  I  P  A  S  W  N  N  H  M  A  T  F  Y  P  L  F  P  N
      1090                      1110                      1130
ATGGTGGTGGCACTTATCCTGAAGTCGTCAATGATTTCCCTTTGAAACTACTTTATTTTA
  G  G  G  T  Y  P  E  V  V  N  D  F  P  L  K  L  L  Y  F  T
      1150                      1170                      1190
CTAATTTAAACTATTTTGTACTGATGTAGCCCTGAGGTAGTTCATGAAAAATGCTGTGCA
  N  L  N  Y  F  V  L  M  *
      1210                      1230                      1250
CTCATTCATGGGAATAAATGTTGGGAAAGCTGAAAAAAAAAAAAAAAAAAAAA

```

0925502 " 010699

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: Moore et al.

Application Serial No.: Unassigned

Art Unit: Unassigned

Filed: Herewith

Examiner: Unassigned

For: **Human FK506 Binding Proteins** Attorney Docket No.: **PF392**

STATEMENT UNDER 37 C.F.R. 1.821 and 1.825

Assistant Commissioner
For Patents
Washington, D.C. 20231

Sir:

Applicants hereby certify under 37 C.F.R. §§1.821(f) and 1.825(b) that the hard copy of the sequence listing and the computer-readable form of such sequence listing submitted herewith, are identical.

Respectfully submitted,

Date: 1/06/99


A. Anders Brookes (Reg. No. 36,373)
Attorney for Applicant

Human Genome Sciences, Inc.
9410 Key West Avenue
Rockville, MD 20850
Phone (301) 610-5761
Fax (301)309-8439

Enclosure(s)
AAB/mbp

09225502 010699

DECLARATION FOR PATENT APPLICATION

As a below named inventor, I declare that I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

Human FK506 Binding Proteins

the specification of which is being filed concurrently herewith.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in 37 C.F.R. § 1.56.

I hereby claim foreign priority benefits under 35 U.S.C. § 119(a)-(d) or § 365(b) of any foreign application(s) for patent or inventor's certificate, or § 365(a) of any PCT international application, which designated at least one country other than the United States listed below, and have also identified below any foreign application for patent or inventor's certificate, or PCT international application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application(s):

Priority Claimed
Yes No

(Number) (Country) (Day/Month/Year Filed)

I hereby claim the benefit under 35 U.S.C. § 119(e) of any United States provisional application(s) listed below.

(Application Serial No.) (Filing Date)

I hereby claim the benefit under 35 U.S.C. § 120 of any United States application(s), or under § 365(b) of any PCT international application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by the first paragraph of 35 U.S.C. § 112, I acknowledge the duty to disclose information that is material to patentability as defined in 37 C.F.R. § 1.56 that became available between the filing date of the prior application and the national or PCT international filing date of this application.

(Application Serial No.) (Filing Date) (Status: patented, pending, abandoned)

I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith: A. Anders Brookes (Reg. No. 36,373), James H. Davis (Reg. No. 40,582) and Kenley Hoover (Reg. No. 40,302) of Human Genome Sciences, Inc. 9410 Key West Avenue, Rockville, Maryland, 20878.

09225502 010699

Full name of first joint inventor: Paul A. Moore

Full name of first joint inventor: Ying-Fei Wei

Full name of additional joint inventor: Steven M. Ruben

Inventor's signature: _____ Date: _____
 Residence: 18528 Heritage Hills Drive, Olney, MD 20832 Citizenship: US
 Post Office Address: same as above